CARLETON UNIVERSITY

August 2014

Final

EXAMINATION

August 2014

DURATION: 3 HOURS No. of Students

Department Name & Course Number: Mathematics and Statistics STAT3502A Final

Course Instructor(s) Dr. Z. Montazeri

AUTHORIZED MEMORANDA

NON-PROGRAMMABLE CALCULATORS ARE ALLOWED

Students **MUST** count the number of pages in this examination question paper **before** beginning to write, and report any discrepancy immediately to a proctor. This question paper has **14** pages, page 16 is blank and is to be used for rough calculation. In addition to this question paper, students don't require an examination booklet.

This examination question paper **MAY NOT** be taken from the examination room.

This examination question paper **MAY NOT** be released to the library.

Last Name : ____________________ **First Name :** ____________________

Student Number : ____________

Problem	Maximum Mark	
MC	30	
1	13	
2	7	
3	4	
4	7	
5	6	
6	5	
7	4	
8	6	
9	6	
10	5	
11	7	
Total	100	

PART I: Multiple-choice questions (choose one answer ONLY)

Each multiple-choice question has 3 marks.

1. In order to test $H_0 : \mu = 50$ vs $H_a : \mu \neq 50$, a random sample of 9 observations (from a normally distributed population) is obtained, yielding $\bar{x} = 61$ and $s = 21$. What is the P-value of the test?

   (a) greater than 0.1 **(b)** between 0.05 and 0.10

   (c) between 0.01 and 0.05 **(d)** less than 0.01.

2. Suppose we wish to test the null hypothesis that a new drug is not effective in reducing blood pressure versus the alternative that it does reduce blood pressure. A ———— error would be made by concluding that the new drug is ————- in blood pressure reduction if in fact the new drug is ———— in blood pressure reduction.

   (a) type I; not effective; not effective **(b)** type I; not effective; effective

   (c) type II; effective; not effective **(d)** type II; not effective; effective

3. Let X_1 and X_2 be independent normal random variables with means $\mu = 3$ and variances $\sigma^2 = 4.5$. Let $T = X_1 + X_2$. Then $P(T \leq 12)$ is

   (a) 0.9772 **(b)** 0.0548 **(c)** 0.8757 **(d)** 0.9452

4. Which of the following is a consequence of the Central Limit Theorem?

   (a) A large population will be normally distributed with mean μ and variance σ^2/n

   (b) A large sample will be normally distributed with mean μ and variance σ^2/n

   (c) The sample mean $\bar{X}$ is approximately normally distributed with mean μ and variance σ^2/n for large sample

   (d) The sample mean $\bar{X}$ is approximately binomially distributed with mean μ and variance σ^2/n

5. X and Y are independent random variables with $V(X) = V(Y) = 4$. Let $T = X - Y$. What is $Corr(X, T)$?

   vspace1mm

   (a) 0 **(b)** $\frac{3}{2\sqrt{2}}$ **(c)** $\frac{1}{4\sqrt{2}}$ **(d)** $\frac{1}{\sqrt{2}}$

6. After taking 90 observations, you construct a 90% confidence interval for μ. You are told that your interval is 3 times too wide (i.e., your interval is 3 times wider than what was required). Your sample size should have been

   (a) 30 **(b)** 270 **(c)** 810 **(d)** 10

7. Suppose we repeat an experiment identically and independently 100 times. Each time we construct a 99% confidence interval for μ via the t-distribution. Let X = the number of times the confidence interval fails to contain the true value of μ. The distribution of X is

   (a) Normal with μ=99 and σ^2=0.99 **(b)** Normal with μ=0 and σ^2=1
   (c) Binomial with n=100 and p=0.99 **(d)** Binomial with n=100 and p=0.01

8. In a histogram, the proportion of the total area which must be to the right of the mean is

   (a) Less than 0.50 if the histogram is skewed to the left
   (b) Exactly 0.50 (always)
   (c) Exactly 0.50 if the histogram is symmetric and unimodal
   (d) More than 0.50 if the histogram is skewed to the right

9. The number of calls received at a fire station can be described according to a Poisson distribution with a mean of 2 calls per day. What is the probability that over a 5 day period there will be at least one call?

   (a) $\frac{2e^{-2}}{2}$ **(b)** $1 - \frac{2e^{-2}}{2}$ **(c)** $1 - e^{-5}$ **(d)** $1 - e^{-10}$

10. An oil firm plans to drill 20 wells, each having probability 0.2 of striking oil. Each well costs \$20,000 to drill; a well which strikes oil will bring \$750,000 in revenue. Find the expected gain from the 20 wells.

    (a) \$15,000,000 **(b)** \$3,000,000 **(c)** \$2,600,000 **(d)** \$1,500,000

PART II: For the following questions clearly show all of your work

1. Assume that head sizes (circumference) of new recruits in the Canadian armed forces can be approximated by a normal distribution with a mean of 22.8 inches and a standard deviation of 1.1 inches.

   a. [2] What proportion of recruits have head sizes between 22 and 23 inches?

   b. [3] Find the value of C such that 5% of the head sizes exceed C inches?

   c. [3] What is the probability that the average head size of the next 49 recruits will be less than 22.5 inches?

   d. [2] If the distribution of head sizes was not normal, would your calculations in part (c) still make sense? Why?

   e. [3] What is the probability that at least 10 of the next 30 recruits will have head sizes between 22 and 23 inches?

2. Suppose that the lifetime of an electronic component follows an exponential distribution with parameter λ.

   a. [3] If the probability that the lifetime is less than one is 0.095, what is the value of λ?

   b. [2] By using part (a), find the probability that the lifetime is less than 10.

   c. [2] By using part (a), find t such that the probability of lifetime greater than t is 0.01.

3. [4] A study conducted by a commuter train transportation authority involved surveying a random sample of 200 passengers. The results show that a customer had to wait on the average 9.3 minutes with a standard deviation of 6.2 minutes to buy his or her ticket. Construct a 95% confidence interval for μ, the true mean waiting time.

4. [7] Consider the random sample $X_1, X_2, \cdots, X_n$ from the pdf

$$f(x;\theta) = 0.5(1+\theta x), \quad -1 \leqslant x \leqslant 1, \quad -1 \leqslant \theta \leqslant 1.$$

Show that $\hat{\theta} = 3\bar{X}$ is an unbiased estimator of θ. (Hint: How are $E(X)$ and $E(\bar{X})$ related?)

5. [6] The joint distribution of X and Y is given by the following joint probability density function (pdf)

$$f(x,y) = \begin{cases} \frac{1}{8}(x^2 - y^2)e^{-x} & \text{if } 0 \leqslant x < \infty, \text{ and } -x \leqslant y \leqslant x \\ 0 & \text{otherwise.} \end{cases}$$

Compute the marginal pdf of X, and the conditional pdf of Y given $X = 1$.

6. [5] Let $X_1, X_2, \cdots, X_n$ represent a random sample from the following distribution

$$f(x;\theta) = \frac{x}{\theta} e^{\frac{-x^2}{2\theta}}, \quad x > 0$$

Find the maximum likelihood estimate for θ.

7. Consider the following situation:

$A =$Visa Card, $B =$Master Card, $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \cap B) = 0.25$.

a. [2] Find $P(B'|A)$?

b. [2] Given that an individual is selected at random and that he or she has at least one card, what is the probability that he or she has a Visa card?

8. In order to estimate the difference in verbal SAT scores for high school students intending to major in engineering and those intending to major in language and literature, independent random samples of size 15 were obtained from the two groups. The results were as follows:

Engineering	$\overline{Y}_1 = 446$	$s_1 = 42$
Language/literature	$\overline{Y}_2 = 534$	$s_2 = 45$

a. [5] Find a 95% confidence interval estimate for the difference in verbal SAT scores between students intending to major in engineering and those intending to major in language/literature. Be sure to state all assumptions required.

b. [1] What is the interpretation of the confidence interval in part (a) ?

9. [6] Suppose that, based on past experience, it is known that a lie detector test will indicate that an innocent person is guilty with probability .08, while the test will indicate that a guilty person is innocent with probability .15. Suppose further that 10% of the population under study has committed a traffic violation. If a lie detector test indicates that a randomly chosen individual form this population has committed a traffic violation, what is the probability that this person is innocent of committing a traffic violation?

10. To test the ability of auto mechanics to identify simple engine problem, an automobile with a single such problem was taken in turn to 72 different car repair facilities. Only 42 of the 72 mechanics who worked on the car correctly identified the problem.

    a. [2] Does this strongly indicate that the true proportion of mechanics who could identify this problem is less than 0.75?($\alpha = 0.05$)

    b. [2] Calculate an upper prediction bound for population proportion using a confidence level of 95%.

    c. [1] Does the result of part (b) confirm part (a)?

11. A sample of 12 radon detectors of a certain type was selected, and each was exposed to 100 pCi/L of radon. The resulting readings were as follows

    104.3 89.6 89.9 95.6 95.2 90.0 98.8 103.7 98.3 106.4 102.0 91.1

    a. [4] Does this data suggest that the population mean reading under these conditions differs from 100? State and test the appropriate hypotheses using $\alpha = 0.05$

    b. [3] Construct a 99% confidence interval for the population mean reading.

    $\sum x_i = 1164.9, \quad \sum x_i^2 = 113493.25$

Formulae Sheet

$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n-1}$

$E(X) = \mu = \sum_x xp(x) = \int xf(x)dx,$

$\sigma^2 = \sum_x (x-\mu)^2 p(x) = \sum_x x^2 p(x) - \mu^2 = \int_x x^2 f(x)dx - \mu^2$

If $A_1, A_2, \cdots, A_k$ are mutually exclusive and exhaustive events. Then for an event B,

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_k)P(A_k) = \sum_{i=1}^{k} P(B|A_i)P(A_i)$$

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)} \quad j = 1, \cdots, k$$

$f_X(x) = \sum_y P(x,y) = \int_y f(x,y)dy, \qquad f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

$E(XY) = \int_x \int_y xyf(x,y)dydx, \qquad \text{Cov}(X,Y) = E(XY) - \mu_X\mu_Y, \quad \rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$

If $E(X) = \mu$ and $V(X) = \sigma^2$, then $E(\bar{X}) = \mu; \quad V(\bar{X}) = \sigma^2/n; \quad Z = \frac{X - E(X)}{\sqrt{V(X)}}$

$V(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Cov}(X_i, X_j)$

Distributions

- Binomial: $P(X = x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \qquad \mu = np, \qquad \sigma^2 = np(1-p)$
- Hypergeometric $P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \quad \mu = n\frac{M}{N}, \quad \sigma^2 = n\left(\frac{M}{N}\right)\left(1 - \frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$
- Poisson $\quad P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad E(X) = \lambda = \text{Var}(X)$
- Gamma $f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}, \quad x \geqslant 0 \qquad E(X) = \alpha\beta, \quad V(X) = \alpha\beta^2$
- Exponential $f(x) = \lambda e^{-\lambda x} \quad x \geqslant 0 \quad E(X) = 1/\lambda, V(X) = 1/\lambda^2$
- Uniform $f(x) = \frac{1}{b-a} \quad a \leqslant x \leqslant b \quad E(X) = (a+b)/2, V(X) = (a+b)^2/12$

Important Test Statistics	**Confidence Interval**
$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	$\left(\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$
$\frac{\overline{X}-\mu_0}{s/\sqrt{n}}$	$\left(\overline{X} \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right)$
$\frac{\overline{X_1}-\overline{X_2}-D_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	$\left(\overline{X_1} - \overline{X_2} \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$
$\frac{\overline{X_1}-\overline{X_2}-D_0}{\sqrt{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)}}$	$\left(\overline{X_1} - \overline{X_2} \pm t_{n_1+n_2-2,\alpha/2}\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}\right)$
$\frac{\hat{p}-p_0}{\sqrt{\frac{p_0 q_0}{n}}}$	$\left(\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$
$\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{pq}{n_1}+\frac{pq}{n_2}}}$	$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$

a

a

a